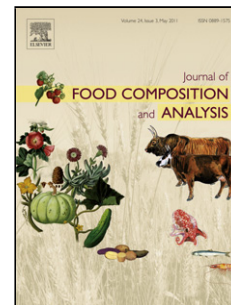


Accepted Manuscript

Title: A procedure for grouping food consumption data for use in food allergen risk assessment

Authors: Sophie Birot, Charlotte B. Madsen, Astrid G. Kruizinga, Tue Christensen, Amélie Crépet, Per B. Brockhoff



PII: S0889-1575(17)30018-2
DOI: <http://dx.doi.org/doi:10.1016/j.jfca.2017.01.008>
Reference: YJFCA 2822

To appear in:

Received date: 26-8-2016
Revised date: 14-12-2016
Accepted date: 11-1-2017

Please cite this article as: Birot, Sophie., Madsen, Charlotte B., Kruizinga, Astrid G., Christensen, Tue., Crépet, Amélie., & Brockhoff, Per B., A procedure for grouping food consumption data for use in food allergen risk assessment. *Journal of Food Composition and Analysis* <http://dx.doi.org/10.1016/j.jfca.2017.01.008>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Original Research Article

Title: A procedure for grouping food consumption data for use in food allergen risk assessment

Authors

Sophie BIROT^a, Charlotte B. Madsen^b, Astrid G Kruizinga^c, Tue Christensen^b, Amélie Crépet^d, Per B. Brockhoff^a

Authors' affiliations

^a DTU Compute, Richard Petersens Plads, DK-2800 Kgs. Lyngby, Denmark sobi@dtu.dk

^b National Food Institute, Technical University of Denmark, Denmark

^c The Netherlands Organization for Applied Scientific Research (TNO), Zeist, The Netherlands

^d ANSES, French Agency for Food, Environmental and Occupational Health Safety, 14 rue Pierre et Marie Curie, 94701 Maisons-Alfort, France.

Highlights

- Customized clustering method followed by allergen risk assessment per food group
- Food groups automatically created for France, Netherlands and Denmark
- Danish consumption data available per food group created with the procedure

Abstract

Food allergic subjects need to avoid the allergenic food that triggers their allergy. However, foods can also contain unintended allergens. Food manufacturers or authorities need to perform a risk assessment to be able to decide if unintended allergen presence constitutes a risk to food allergic consumers. One of the input

parameters in risk assessment is the amount of a given food consumed in a meal. There has been little emphasis on how food consumption data can be used in food allergen risk assessment. The aim of the study was to organize the complex datasets from National Food Consumption Surveys from different countries (France, Netherlands and Denmark) to be manageable in food allergen risk assessment. To do this, a two-step method was developed. First, based on initial groups of similar food items, the homogeneity of consumption was evaluated using a customized clustering method. Then, the risk was calculated for each initial food group and its subgroups to verify if it also represents a relevant difference in risk. Forty-eight food groups were designated in Denmark (53 in the Netherlands, 54 in France). Finally, summary statistics and names for each food group for the Danish data illustrate the results when applying the procedure.

Key words

Food allergy; National Food Consumption Surveys; food groups; probabilistic risk assessment; food analysis; food composition; automatic procedure; allergen contamination

1. Introduction

In Europe the lifetime prevalence and the point prevalence of self-reported food allergy is around 17% and around 6%, respectively. These figures probably reflect an overestimation of the true prevalence that also varies between countries, but the true prevalence is not known (Nwaru et al., 2014). In food allergic subjects exposure to food allergens can trigger an acute allergic reaction, which can have a wide range of symptoms: from itching in the mouth to life-threatening anaphylactic shocks (Bock et al., 2001). For this reason allergic consumers need to avoid consumption of the food containing the harmful allergen (Fernández-Rivas and Asero, 2014). This can be done by avoiding foods where the allergen is in the list of ingredients, as the

European Directive 2003/89/EC makes mandatory the labelling of all ingredients derived from priority allergenic sources. However unintended contamination during the production may occur and as the allergen is not part of the ingredients, it is not mentioned in the ingredient list. This unintended contamination could be due to shared production lines or equipment, also in the storage of ingredients. Thus, products with unintended allergens are potentially harmful for the allergic consumers. The increasing use of “may contain” labelling could help allergic consumers manage their food allergy. However, the lack of regulation in Europe has led to the over-use of such labels. The labels are actually used even though many products are not contaminated (DunnGalvin et al., 2015).

To be able to decide if the unintended allergen traces constitute a risk to food allergic consumers, it is necessary to perform a risk assessment. There are several ways of performing risk assessment for the unintended presence of allergens in food but the underlying need for data is similar (Madsen et al., 2009). Three types of data are needed: The concentration level of the allergen in the food in question, data on the dose that may elicit an allergic reaction within the population allergic to that food and the estimated consumption of the contaminated food. The dose that may elicit a reaction comes from clinical studies where food allergic subjects are challenged with the allergenic food (Taylor et al., 2014). The food producer or a control authority typically provides the level of concentration of the allergen in the food product.

At the moment in food allergen risk assessment, National Food Consumption Surveys are used to estimate consumed quantities for a wide range of food products. However, those surveys are primarily designed to answer nutritional questions, are not performed in a specific food allergic group and the level of detail may not be suitable for the need of food allergen risk assessment. In a peanut allergic population, initial investigations on consumption were performed in the MIRABEL project (Crépet et al., 2015). However, this has limited value for the allergen risk assessment as the food record was performed for limited food groups

and consumption estimates were not recorded per eating occasion. Moreover, there has been little investigation on how food consumption data can be best used in food allergen risk assessment.

Furthermore, risk assessments are also usually performed to assess the risk of various chemicals in the toxicology field. Consumption of different food products are also used to estimate the risk in different populations. The way of grouping the food products will depend on the chemical for which this risk assessment is done. For example, when there is an interest in the risk of high doses of a vitamin, foods might be grouped according to the specific vitamin content. However, the concentration of unintended allergen presence cannot be a criterion, since the level will vary in foods according to the production site and batch. Therefore, for food allergen risk assessment food groups are needed according to the consumption pattern.

The aim of the work presented in this paper was to develop a standardized, automated procedure to organise data from national food consumption surveys to be suitable to use in food allergen risk assessment. Building on the previous work done by TNO, food groups were proposed using a two-step procedure and national Food Consumption Surveys from three countries were made available within the iFAAM (Integrated Approaches to Food Allergen and Allergy Management) project: Denmark, France and Netherlands. The procedure was tested on consumption data from three different countries, to investigate if the results were similar and stable across countries. Thus, the focus of the paper is on grouping food products within similar commodities that also have the same consumption patterns, to ease the choice of the correct food group. The allergen risk assessment can then be performed with a food category which has the correct consumption distributions regarding the unintended contaminated food product.

2. Materials and Methods

In this study we built on the TNO food grouping approach to organize the large amount of food consumption data. Thus, the first step was to organize the food consumption data from the three countries into groups

comparable to the original groups from TNO, based on Dutch 2003 consumption data. An automated two-step procedure was developed building on the unpublished work from TNO describing an approach for food grouping for allergen risk assessment (Figure 1). First, during the clustering method step, the initial groups were automatically split using a common statistics criterion. The aim of the food grouping is to have logical food groups with homogeneous consumption patterns and similar risk outcome. I.e., food groups that gather food items with similar consumption pattern that could be interchangeable to estimate the risk of allergic reaction. Secondly, the consequences of splitting the initial food groups from the National Food Consumption Survey from each country were investigated. Based on a fixed criterion, we checked if the subdivisions suggested by the clustering method were needed or the difference between risks was too small to make a split meaningful.

2.1. National Food Consumption Surveys (Netherlands, France and Denmark)

Detailed food consumption data from the most recent National Food Consumption Survey in Netherlands (Van Rossum et al., 2011), France (Dubuisson et al., 2010) and Denmark (Pedersen et al., 2008) were used. Food consumption in the Netherlands was recorded with two non-consecutive 24 hour recalls. In France and Denmark, a pre-coded seven days record was used. We used data from participants aged 18 to 75 years from all three surveys. The food consumption was recorded for each eating occasion, defined as breakfast, lunch, dinner and other distinct eating occasions during the day. The maximum consumption on a single eating occasion for each consumer and food item was used. As food allergy is an acute reaction, where the allergen if not causing reaction, will not be accumulated, the relevant data are consumption at the meal level (Madsen et al., 2014). All three countries use a common coding system for food items: the Food Ex 2 coding developed by EFSA (European Food Safety Authority, 2015).

2.2. TNO food grouping approach and initial food grouping

TNO has performed allergen risk assessments for several years (Kruizinga et al., 2008; Spanjersberg et al., 2007). Data from the Dutch National Food Consumption Survey 2003 (Ocke et al., 2005) were organised into

food groups. The basis of the original grouping was the consumption survey coded according to the Dutch food composition table. First a selection was made of the processed foods based on the assumption that contamination with allergens is not present in raw foods, when products are not processed it is assumed that allergen cross-contamination will not occur. Furthermore in the original approach composite dishes were split into their original foods (for example a rice dish was split into rice, meat and vegetables). This was only done for composite dishes that form a complete meal. Based on a standard food grouping used for the Dutch food composition table, consumption data, portion size and if needed, expert judgement, food products were combined into 1 groups applicable for food allergen risk assessment e.g. cookies, ice cream, bread. For example the food group potatoes in the Dutch food composition table was split into potato powder and potato products, since there is a large difference in the amount consumed. In the end this resulted in 42 food groups with similar food items and similar intakes, which were validated using allergen risk assessment. These 42 food groups were applied to the food consumption data from the three countries coded using the FoodEx2 categorization, resulting in the initial food grouping used as a basis for this study. FoodEx2 is a standardized food categorization and description system designed by EFSA with various levels of detail. For this study the most detailed information was used.

2.3. Clustering method

The first analysis was to assess the differences between food item consumption distributions within each of the initial groups for each country. This was done on the consumption quantities of the three countries separately. The results of these analyses give an indication on how the initial groups can be subdivided to new food groups with similar consumption patterns.

2.3.1. Differences between food items consumption distribution

The two non-parametric tests, the Kolmogorov-Smirnov (KS) test and the Cramér-van Mises test, commonly used to check if two distributions are drawn from the same distribution, were used.

Comparing the probability distributions of two food item consumed quantities, the Kolmogorov-Smirnov statistic can be written as follow (Conover, 1971):

$$D_{N, M} = \sup_x |F_N(x) - G_M(x)|$$

where N and M are the numbers of observations for the empirical distribution functions F_N and G_M of the first and second food items. And sup is the supremum function (maximum).

The Cramér-van Mises criterion is defined as (Anderson, 1962):

$$T = \int_{-\infty}^{\infty} [F_N(x) - G_M(x)]^2 dH_{N+M}(x)$$

where N and M are the numbers of observations for the empirical distribution functions, F_N and G_M of the food items N and M. And, $H_{N+M}(x)$ is the empirical distribution function of the 2 food items together.

The KS and Cramér test statistics measure the distance between the food consumption distributions in different ways. The KS test statistics evaluates the difference between distributions based on one point only (the maximum difference, represented by the arrow in Figure 1) whereas the Cramér distance calculates the difference using the area between the two distributions' curves (grey part in Figure 1).

2.3.2. Distances matrices within each initial food group

Therefore, we calculated the KS and Cramer test statistics for each pair of food items. Two food items with similar distribution will have a small test statistics. And oppositely, the difference in distribution between two food items will increase with the test statistics value. Thus, the test statistics quantifies the distance between the empirical distribution function of two food items and measures the similarities between food items within each initial group. Thus, we construct two matrices distances with the KS and Cramér test statistics calculated for each pair of food items within each food group in order to quantify the difference between each food item within each initial group.

2.3.3. Automatic clustering method

Then, the distance matrices are used as an input to the clustering method. It resulted in different suggestions for subdivisions; from not splitting an initial food group to having one food item in each subdivision. The complete-linkage clustering method is used, so the number of food items in each suggested subdivision is balanced (Sørensen, 1948). Finally, the optimal split is automatically selected based on the average silhouette index (Rousseeuw, 1987). For each subdivision, this index measures how food items consumption patterns within a subdivision are similar to each other within a subdivision compared to the ones within the other subdivisions. Thus, this index assesses which food items lie together within each subdivision and which food items could be in another subdivision. So, the validity and the quality of the different suggested subdivisions within each initial food group is evaluated and only the subdivisions for which the food items have the most similar consumption patterns are used for the next step of the procedure.

2.4. Risk simulation and clustering validation

Once the optimal split, as described above, was selected the next step was to investigate if the proposed subgrouping is appropriate in terms of estimated risk of an unexpected allergic reaction. This was done by comparing the risk in the subgroups with the risk in the group before subdivision. To calculate the risk a probabilistic risk assessment method was used: the method uses challenge data with allergenic foods, contamination data and consumption data. From these three inputs, the percentage of the population likely to have an allergic reaction is estimated, i.e. the chance for an allergic consumer to eat a high enough amount of allergen to trigger a reaction (Rimbaud et al., 2010; Spanjersberg et al., 2007).

2.4.1. Oral food challenge

Food allergen risk assessments use data where food allergic subjects are challenged in a double blind placebo controlled food challenge (DBPCFC) with the offending food, to determine the dose that elicits an objective allergic reaction. If there is enough data available it is possible to create distributions (Taylor et al., 2014). In this study, a selection described in Taylor et al. (2014) was used, this selection was based on already published data. The threshold distributions are based on discrete NOAEL (No Observable Adverse Effect

Level) and LOAEL (Lowest Observable Adverse Effect Level) values from 158 DBPCFCs for peanut (Anagnostou et al., 2009; Atkins et al., 1985; Blumchen et al., 2010; Clark and Ewan, 2008; Hourihane et al., 1997; Leung et al., 2003; Lewis et al., 2005; Nelson et al., 1997; Nicolaou et al., 2010; Oppenheimer et al., 1992; Patriarca et al., 2006; Wainstein et al., 2010) and 43 for soy (Ballmer-Weber et al., 2007; Fiocchi et al., 2003; Magnolfi et al., 1996; Zeiger et al., 1999). These two allergens were selected because they represent a large difference in challenge data. In order to assess the homogeneity of consumption patterns, the challenge data are modelled using one of the distributions commonly used to fit the challenge data in the allergen risk assessment field (Crevel et al., 2007). Thus, data were fitted using the Weibull distribution, a distribution conventionally used to fit allergy thresholds data (Crevel et al., 2007). The Survival package (version 2.38.3) from R software (version 3.3.0) was used to fit the data (R Core Team, 2015). The fitted distribution was used to calculate the risk of an allergic reaction for the selected food item's distribution. These two threshold distributions are used as a model computational framework for all the food groups and items.

2.4.2. Allergen concentration in food

A range of theoretical concentration data (1, 10, 100 and 1000 ppm protein) were selected to assess the change in risk with increasing contamination. As the contamination happens unintentionally during the production, the level of concentration is not known and the range of concentration helps to evaluate the impact of different level of concentration on the outcome of the risk assessment. This allowed us to evaluate the behaviour and the homogeneity of different quantities of food consumed according to different concentration levels of peanut protein and soy protein.

2.4.3. Risk simulation

As the aim of the analysis was to evaluate the impact of grouping the food items in different food groups according to consumption, the probabilistic risk assessment was performed in a fixed framework of contamination and challenge data. It is assumed that 100% of the products are contaminated and 100% of

the consumers were allergic. The simulations were iterated $K = 1000$ times for $N = 10,000$ re-samplings among consumers for each country separately. Thus, for each iteration K , first N consumers were sampled with replacement according to consumers' sampling weights in the surveys. Then, one eating occasion of a food item among the ones consumed in a food group was randomly sampled. This distribution was used in the risk calculation to represent the consumption in a food group. Regarding the threshold value, for each iteration K , and each consumer N , a threshold was sampled from the Weibull model presented in the previous section. Finally, the consumption value was multiplied with each unintended allergen level from the range 1, 10, 100 and 1000 ppm.

Thus for each iteration K and each consumer N , we compared the amount of allergens consumed (level of contamination $\times$ amount consumed) to the simulated thresholds. So for each iteration K , we calculated the percent of allergic reaction among the N consumers. Therefore, for each group and each subgroup, eight different risks were calculated corresponding to the four levels of contamination and the two allergens.

2.4.4 Maximum absolute difference and decision criteria

To assess the impact of difference in consumption on the outcome of the probabilistic risk assessment, the maximum relative difference between the risk of each proposed subgroup and the risk of each initial food group was calculated:

$$\max \text{ relative difference Group} = \max_{\text{Subgroup1}, \text{Subgroup2}} \left(\frac{|Risk_{\text{Subgroup}} - Risk_{\text{Group}}|}{Risk_{\text{Group}}} \right)$$

In order to decide whether a group needed to be divided, a formal decision criterion was useful. It helped to assess if the difference between the overall group risk and the subgroup risks is high. Based on similar investigation, no criterion was defined before, thus we chose to develop an objective decision criterion based on the actual data. Various summary statistics (mean, median, different quantiles) for the maximum relative differences of the risk across the contamination levels and allergens were calculated in order to evaluate the

magnitude of the maximum relative and absolute differences across the initial food groups. This short study helped to make a decision on what is a high difference in risk between a group and its suggested subdivisions. The median was the only summary statistics that was stable across the conditions (allergen $\times$ level of concentration) introduced to assess the differences in risk of allergic reaction between an initial group and its suggested subdivisions. Thus, based on those investigations, the median of the maximum relative difference on the risk simulated across all the contamination levels, allergen and groups was seen as the best choice to decide if the original groups need to be split or not. Therefore, the criteria were calculated for each country. Groups with a low number of food items, with no subdivisions suggested, were removed from the median calculation. For each group, the criterion was used for all concentration levels and the two allergens. Therefore, one initial group is automatically split if the median of the maximum relative difference in risk of the four concentration levels and the two allergens were all above the threshold. Thus, the criterion can be interpreted as such: when the risk connected to consumption of the different subgroups of an initial group is calculated, we can then calculate the maximum relative difference in risk between the overall group's risk and the subgroups' risks. We can then compare them to the criterion to evaluate whether the risks difference between subgroups is so high that we cannot use the whole group consumption distribution to calculate the group risk.

3. Results

3.1. Initial grouping of National Food Consumption Surveys

Table 1 displays the numbers of consumers, the number of food items in the national food consumption surveys and the number of initial groupings. The difference in number of food items reflects the difference in the level of detail when coding foods at the national level.

3.2. Criterion for each country and example of using the criterion for groups design

The median of the maximum relative difference is 15.1% (Denmark), 13.6% (France) and 16.3% (The Netherlands). This figure for each country was used as the country specific criterion.

The breakfast products' group in Denmark is used to illustrate the grouping design and the application of the method (Table 2). The maximum difference for all the contamination levels and the 2 allergens are higher than the threshold (15.1%). The differences between subgroups are high and this group needs to be split, so single food items need to be allocated to subgroups with more homogeneous consumption patterns.

3.3. Groups outline based on clustering and risk analysis

The first split of the initial groups was based on the clustering method described in the section 2.3. Based on these results, each initial group should be divided into one, two or more subgroups. To test the relevance of this grouping to risk assessment, the subgroups were evaluated using the risk calculation. Table 3 shows the need for subdivisions based on clustering and the reduced need for subdivision if the analysis is based on risk using the decision criterion. The procedure was performed on the three countries separately.

It was not possible to integrate a criterion to assess the impact of the difference in consumption on the risk during the clustering step. It can be seen that using the risk to evaluate the need of subdividing the initial groups reduces the number of groups that need to be subdivided for the three countries. The procedure minimizes the number of groups with homogeneous consumption patterns.

3.4. Examples with Danish data

Groups of breakfast products, meat, cheese and milk products from the Danish data were chosen to illustrate the procedure. For each group, the risk of an allergic reaction was plotted to illustrate the difference in risk between the subgroups across the concentration levels and for the two allergens. In addition, in order to assess the variability within each group and subgroup, the coefficient of variation was calculated ($CV = \frac{\text{standard deviation}}{\text{mean}}$). It measures the dispersion of the consumption around the mean. Thus, the CV evaluates the change in homogeneity of consumption patterns between keeping or splitting an initial group.

Those numbers are reported in a table for each group. Finally, the maximum relative difference for the contamination levels and allergens was also described in a table when it supports the explanation of the choice of splitting an initial group or not.

3.4.1. Splitting the initial group: breakfast products and meat products

Breakfast products were divided into two subgroups: breakfast products eaten unprocessed (subgroup 1) and porridges (subgroup 2). There is a large difference in risk between the two subgroups (see Table 2 and Figure 3). The maximum relative differences between the risk of the breakfast products group and its two subgroups for each contamination level and each allergen are all above 15.1%; hence it shows that the difference in consumption between the two subgroups is large, according to the criteria used. Subgroup 2 has a higher consumption (hence risk) than the subgroup 1, so if the combined consumption is used, it will result in an underestimation of the risk for the subgroup 2 and an overestimation for subgroup 1. Moreover, in Table 4, the coefficient of variation shows that subdividing the breakfast product group also decreases the variability within the consumption subgroups. Thus, the breakfast group is split into two subgroups, so groups with more homogeneous consumption patterns can be used as an input to the risk assessment.

The meat product group was also divided into two. Again, a large difference in risk between the two subgroups can be seen in Figure 4. In a similar way, dividing this group will help the user to use a food distribution with a lower variability for the risk assessment (Table 5).

3.4.2. Keeping the initial group: cheese and milk and milk products

Figure 5 and Table 7 shows that the difference in risk between the two cheese subgroups is very small (all the maximum relative differences are under 15.6%). For this reason the cheese group was not subdivided during the procedure. It is confirmed by the mean consumption for the cheese groups and the two subdivisions: the consumption means for the food group and its subdivisions are similar. This is confirmed by calculating the coefficient of variation. For the second subgroup, the coefficient of variation is also smaller when the cheese

group's consumption distribution is used (Table 7). That is why there is no reason in splitting the cheese group and then increasing the complexity of the food groups.

Figure 6 and Table 8 show the group milk products risks and maximum relative differences; the difference in risk between the two subgroups in the milk products group is small for all concentration levels and the two allergens according to the criterion. In Table 9, it can be seen that there is a small difference in consumption in the two subgroups and this difference does not lead to a large difference in risk (according to the criterion). So, the milk and milk products example illustrates the tradeoff between first splitting the group because of the difference in consumption between subgroups within the initial food group based on clustering and the fact that the difference is not high enough according to the criterion to split the initial food group. Moreover, the coefficients of variation for the milk product group and its two subgroups are all quite similar, so the homogeneity of consumption patterns will not be improved by splitting this group. Therefore, using the consumption of the whole milk products group in the risk assessment will be adequate to describe its consumption.

4. Discussion/Conclusion

In this paper, we present a method to organise food consumption data in groups for use in food allergen risk assessment. The food consumption data from Danish, Dutch and French surveys were initially grouped, building on the food grouping approach developed by TNO using the Dutch 2003 survey. The initial groups were validated or split based on the results from the three countries. This resulted in defining 48 food groups in Denmark, 53 in Netherlands and 54 in France. With this study, we provide a method that can be used for any other consumption database from any other country.

In food allergen risk assessment, the relevant scenario for consumption is the meal. The data used may include multiple consumptions of the same food from different days. As the aim was to protect the allergic consumers from an adverse reaction, we chose a conservative approach and used the maximum

consumption of a food item per person on a single eating occasion. Selecting the maximum consumption assumed a slight overestimation of the risk of allergic reaction, as the average consumption is lower than the one used to estimate the risk of allergic reaction. Rimbaud et al. (2010) also tested random selection within the days of the French National Food Consumption Survey and the difference in risk was not found significant. So, we can expect that using the maximum consumption has a relative low impact on the final group design for each country.

The initial food grouping approach developed by TNO could not be directly used for consumption data from other countries because they were based on the previous Dutch National Food Consumption Surveys and Dutch food codes, which limited the applicability for applying the method to other national food databases, and comparing the food consumption patterns of countries. As the amount of data from national consumption surveys are very large, we wanted to develop a method that was more automated based on fixed criteria. However, clustering was based on consumption level and heterogeneity within each initial group, but not comparing the size of the effects across food groups. So, depending on the homogeneity of the initial group, the difference in consumption between subdivisions across initial food groups can be small or high. The second step was designed to investigate the relevance in relation to risk of splitting the groups. This was done by comparing the risk between the original and the new groups formed because of subdivision. During the risk analysis, the criteria choice was made so that no subjective information had influence, so there was no bias in deciding if a difference in risk is high. As no similar work exists about evaluating the magnitude of those differences, other measures than the maximum relative difference between the group split and the subdivisions risks were investigated, like the maximum absolute difference. But, some tests (not detailed here) showed that maximum relative difference was the most stable criterion across the condition to investigate the consequences of subdividing food groups on the risk of allergic reaction. The choice of the median across the allergens and the contamination levels was found to best fit the purpose of having groups with homogeneous consumption patterns with a significant difference in risk.

As highlighted, the homogeneity of consumption patterns was not evaluated across initial groups beforehand. So, some initial groups can have more homogeneous consumption patterns than other initial groups from the beginning. Using the procedure will help to identify which initial group the homogeneity of the consumption patterns could be improved by splitting. The coefficient of variation calculated for some examples illustrates clearly the impact of splitting or not an initial group on the homogeneity of the consumption within a group or subgroups. For example, for the breakfast products group and the meat product group, the variability within the groups newly formed decreased compared to the ones for the initial groups. Furthermore, it was decided to use national criteria, as the criteria are so alike for the three countries that there was no reason to try to develop a common criterion.

The Weibull distribution was used to fit the dose response curve from the challenge data. Crevel et al. (2007) propose three distributions for fitting threshold data: Weibull, log-logistic and log-normal. It is assumed that selecting another threshold distribution would influence the outcome of the risk assessment with a comparable factor, so the results of the food grouping will be similar. Due to high amount of data and the number of comparisons made in the procedure, it was not realistic to investigate the food group design using the three distributions.

By joining the food items because of the reasons mentioned previously, we choose not to perform the risk assessment at the food items level. It is still possible to perform the risk assessment at the food item level, but it was not the aim of the procedure. At the same time, the procedure also increases the number of consumption data points used to assess the risk of allergic reaction, because several food items in a food group can be used to estimate the risk. Therefore, performing a risk assessment for a single food product is often not opportune, since the sample size will be too small.

The fixed framework of concentration levels and allergen threshold distribution controlled the way that the differences in consumption propagated the risk. So, the difference in risks reflected the difference in

consumption, as the concentration levels and thresholds have similar impact on the risk calculation, only the consumption distribution varies from the different groups and subgroups. The whole food consumption distribution was used in the risk assessment. So, the variability in the consumption distribution was taken into account, and the spread of the consumption distribution is expected not to influence the food grouping. Moreover, uncertainty on the parameters of the concentration and threshold distribution can be integrated during the risk calculation (Kruizinga et al., 2008; Rimbaud et al., 2010; Spanjersberg et al., 2007). Thus, the risk could also be estimated with a confidence interval. It was decided not to investigate the impact of those uncertainties, as they were not the focus of this study and it was preferred to evaluate the impact of the difference in consumption on the risk.

The procedure described in the paper is designed in a way that it can be easily applied to any National Food consumption survey in Europe. For each country the number of suggested subdivisions decreases by around 60% after the risk analysis. So, we assume that food groups based on National Food Consumption Surveys from other countries will have a similar number of groups if using the suggested procedure.

Because of similarity in the data among the three countries we have only used data from Denmark in the detailed examples and present the final grouping in Appendix A and B for Denmark only.

In this paper, we present a method that can be used to group food consumption data from one country using the national food consumption survey. In a world where food is exported to many countries it is relevant to be able to do risk assessment on data that cover multiple countries, ideally, e.g., the whole of EU. We have therefore investigated how food consumption data from multiple countries can be combined using data from the three countries. These results will be presented in a later paper.

Acknowledgements

This work was part of the iFAAM project (Integrated Approaches to Food Allergen and Allergy Risk Management, Grant Agreement No. 322147), the national food consumption data for France and Denmark were kindly provided by the work package partners: ANSES and DTU. The national food consumption data of the Netherlands was kindly provided by the National Institute for Public Health and the Environment in the Netherlands. The challenge data for peanut and soy was kindly provided by TNO and FARRP (University of Nebraska – Food Allergy Research and Resource Program). We also thank TNO for providing the information on the food grouping performed on the Dutch data

Bibliography

- Pedersen, A. N., Fagt, S., Groth, M.V., Christensen, T., Biloft-Jensen, A., Matthiessen, J., Andersen, N.L., Kørup, K., Hartkopp, H., Ygil, K.H., Hinsch, H.-J., Saxholt, E., Trolle, E., 2008. Danskernes kostvaner 2003-2008, Danskernes kostvaner 2003-2008.
- Anagnostou, K., Islam, S., King, Y., Deighton, J., Clark, A.T., Ewan, P.W., 2009. British Society for Allergy and Clinical Immunology Annual Conference 2009 Abstracts, in: British Society for Allergy and Clinical Immunology Annual Conference 2009 Abstracts. Blackwell Publishing Ltd, pp. 1937–1958. doi:10.1111/j.1365-2222.2009.03389.x
- Anderson, T.W., 1962. On the Distribution of the Two-Sample Cramer-von Mises Criterion 1148–1159. doi:10.1214/aoms/1177704477
- Atkins, F.M., Steinberg, S.S., Metcalfe, D.D., 1985. Evaluation of immediate adverse reactions to foods in adult patients: II. A detailed analysis of reaction patterns during oral food challenge. *J. Allergy Clin. Immunol.* 75, 356–363. doi:10.1016/0091-6749(85)90072-7
- Ballmer-Weber, B.K., Holzhauser, T., Scibilia, J., Mittag, D., Zisa, G., Ortolani, C., Oesterballe, M., Poulsen, L.K., Vieths, S., Bindslev-Jensen, C., 2007. Clinical characteristics of soybean allergy in Europe: A double-blind, placebo-controlled food challenge study. *J. Allergy Clin. Immunol.* 119, 1489–1496. doi:10.1016/j.jaci.2007.01.049
- Blumchen, K., Ulbricht, H., Staden, U., Dobberstein, K., Beschorner, J., de Oliveira, L.C.L., Shreffler, W.G., Sampson, H.A., Niggemann, B., Wahn, U., Beyer, K., 2010. Oral peanut immunotherapy in children with peanut anaphylaxis. *J. Allergy Clin. Immunol.* 126, 83–91.e1. doi:10.1016/j.jaci.2010.04.030
- Bock, S.A., Muñoz-Furlong, A., Sampson, H.A., 2001. Fatalities due to anaphylactic reactions to foods. *J. Allergy Clin. Immunol.* 107, 191–193. doi:10.1067/mai.2001.112031
- Clark, A.T., Ewan, P.W., 2008. Good prognosis, clinical features, and circumstances of peanut and tree nut

- reactions in children treated by a specialist allergy center. *J. Allergy Clin. Immunol.* 122, 286–289. doi:10.1016/j.jaci.2008.05.015
- Conover, W.J., 1971. Practical Nonparametric Statistics. *Stat.* doi:10.2307/2986830
- Crépet, A., Papadopoulos, A., Elegbede, C.F., Ait-Dahmane, S., Loynet, C., Millet, G., Van Der Brempt, X., Bruyère, O., Marette, S., Moneret-Vautrin, D.A., 2015. Mirabel: An integrated project for risk and cost/benefit analysis of peanut allergy. *Regul. Toxicol. Pharmacol.* 71, 178–183. doi:10.1016/j.yrtph.2014.12.006
- Crevel, R.W.R., Briggs, D., Hefle, S.L., Knulst, A.C., Taylor, S.L., 2007. Hazard characterisation in food allergen risk assessment: the application of statistical approaches and the use of clinical data. *Food Chem. Toxicol.* 45, 691–701. doi:10.1016/j.fct.2006.09.005
- Dubuisson, C., Lioret, S., Touvier, M., Dufour, A., Calamassi-Tran, G., Volatier, J.-L., Lafay, L., 2010. Trends in food and nutritional intakes of French adults from 1999 to 2007: results from the INCA surveys. *Br. J. Nutr.* 103, 1035–48. doi:10.1017/S0007114509992625
- DunnGalvin, A., Chan, C.H., Crevel, R., Grimshaw, K., Poms, R., Schnadt, S., Taylor, S.L., Turner, P., Allen, K.J., Austin, M., Baka, A., Baumert, J.L., Baumgartner, S., Beyer, K., Bucchini, L., Fernández-Rivas, M., Grinter, K., Houben, G.F., Hourihane, J., Kenna, F., Kruizinga, A.G., Lack, G., Madsen, C.B., Clare Mills, E.N., Papadopoulos, N.G., Aldrick, A., Regent, L., Sherlock, R., Wal, J.M., Roberts, G., 2015. Precautionary allergen labelling: Perspectives from key stakeholder groups. *Allergy Eur. J. Allergy Clin. Immunol.* 70, 1039–1051. doi:10.1111/all.12614
- European Food Safety Authority, 2015. The food classification and description system FoodEx 2 (revision 2). EFSA Support. Publ. 8, n/a–n/a. doi:10.2903/sp.efsa.2011.EN-215
- Fernández-Rivas, M., Asero, R., 2014. Risk Management for Food Allergy, Risk Management for Food Allergy. Elsevier. doi:10.1016/B978-0-12-381988-8.00002-6
- Fiocchi, A., Travaini, M., D’Auria, E., Banderali, G., Bernardo, L., Riva, E., 2003. Tolerance to a rice hydrolysate formula in children allergic to cow’s milk and soy. *Clin. Exp. Allergy* 33, 1576–1580. doi:10.1046/j.1365-2222.2003.01781.x
- Hourihane, J.O., Kilburn, S.A., Nordlee, J.A., Hefle, S.L., Taylor, S.L., Warner, J.O., 1997. An evaluation of the sensitivity of subjects with peanut allergy to very low doses of peanut protein: A randomized, double-blind, placebo-controlled food challenge study. *J. Allergy Clin. Immunol.* 100, 596–600. doi:10.1016/S0091-6749(97)70161-1
- Kruizinga, A.G., Briggs, D., Crevel, R.W.R., Knulst, A.C., Bosch, L.M.C. van den, Houben, G.F., 2008. Probabilistic risk assessment model for allergens in food: sensitivity analysis of the minimum eliciting dose and food consumption. *Food Chem. Toxicol.* 46, 1437–1443. doi:10.1016/j.fct.2007.09.109
- Leung, D.Y.M., Sampson, H.A., Yunginger, J.W., Burks, A.W., Schneider, L.C., Wortel, C.H., Davis, F.M., Hyun, J.D., Shanahan, W.R., 2003. Effect of Anti-IgE Therapy in Patients with Peanut Allergy. *N. Engl. J. Med.* 348, 986–993. doi:10.1056/NEJMoa022613
- Lewis, S.A., Grimshaw, K.E.C., Warner, J.O., Hourihane, J.O., 2005. The promiscuity of immunoglobulin E binding to peanut allergens, as determined by Western blotting, correlates with the severity of clinical symptoms. *Clin. Exp. Allergy* 35, 767–773. doi:10.1111/j.1365-2222.2005.02252.x

- Madsen, C.B., Hattersley, S., Buck, J., Gendel, S.M., Houben, G.F., Hourihane, J.O., Mackie, A., Mills, E.N.C., Nørhede, P., Taylor, S.L., Crevel, R.W.R., 2009. Approaches to risk assessment in food allergy: report from a workshop developing a framework for assessing the risk from allergenic foods". *Food Chem. Toxicol.* 47, 480–9. doi:10.1016/j.fct.2008.12.001
- Madsen, C.B., Houben, G., Hattersley, S., Crevel, R.W.R., Remington, B.C., Baumert, J.L., 2014. Risk Management for Food Allergy, *Risk Management for Food Allergy*. Elsevier. doi:10.1016/B978-0-12-381988-8.00006-3
- Magnolfi, C.F., Zani, G., Lacava, L., Patria, M.F., Bardare, M., 1996. Soy Allergy in Atopic Children. *Ann. Allergy, Asthma Immunol.* 77, 197–201. doi:10.1016/S1081-1206(10)63255-3
- NELSON, H., LAHR, J., RULE, R., BOCK, A., LEUNG, D., 1997. Treatment of anaphylactic sensitivity to peanuts by immunotherapy with injections of aqueous peanut extract1. *J. Allergy Clin. Immunol.* 99, 744–751. doi:10.1016/S0091-6749(97)80006-1
- Nicolaou, N., Poorafshar, M., Murray, C., Simpson, A., Winell, H., Kerry, G., Härlin, A., Woodcock, A., Ahlstedt, S., Custovic, A., 2010. Allergy or tolerance in children sensitized to peanut: Prevalence and differentiation using component-resolved diagnostics. *J. Allergy Clin. Immunol.* 125, 191–197.e13. doi:10.1016/j.jaci.2009.10.008
- Nwaru, B.I., Hickstein, L., Panesar, S.S., Roberts, G., Muraro, A., Sheikh, A., 2014. Prevalence of common food allergies in Europe: A systematic review and meta-analysis. *Allergy Eur. J. Allergy Clin. Immunol.* 69, 992–1007. doi:10.1111/all.12423
- Ocke, M., Hulshof, K., van Rossum, C., 2005. The Dutch National Food Consumption Survey - 2003, Methodological issues. *Arch. Public Heal.* 63, 227–241.
- Oppenheimer, J.J., Nelson, H.S., Bock, S.A., Christensen, F., Leung, D.Y.M., 1992. Treatment of peanut allergy with rush immunotherapy. *J. Allergy Clin. Immunol.* 90, 256–262. doi:10.1016/0091-6749(92)90080-L
- Patriarca, G., Nucera, E., Pollastrini, E., De Pasquale, T., Lombardo, C., Buonomo, A., Roncallo, C., Pecora, V., Musumeci, S., Altomonte, G., Alonzi, C., Schiavino, D., Gasbarrini, G., 2006. Oral Rush Desensitization in Peanut Allergy: A Case Report. *Dig. Dis. Sci.* 51, 471–473. doi:10.1007/s10620-006-3157-4
- R Core Team, 2015. R: A Language and Environment for Statistical Computing.
- Rimbaud, L., Heraud, F., La Vieille, S., Leblanc, J.C., Crepet, A., 2010. Quantitative risk assessment relating to adventitious presence of allergens in food: A probabilistic model applied to peanut in chocolate. *Risk Anal.* 30, 7–19. doi:10.1111/j.1539-6924.2009.01322.x
- Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65. doi:10.1016/0377-0427(87)90125-7
- Sørensen, T., 1948. A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Its Application to Analyses of the Vegetation on Danish Commons, *Biologiske Skrifter / Kongelige Danske Videnskabernes Selskab* 5: 1-34
- Spanjersberg, M.Q.I., Kruizinga, a. G., Rennen, M. a J., Houben, G.F., 2007. Risk assessment and food allergy: the probabilistic model applied to allergens. *Food Chem. Toxicol.* 45, 49–54. doi:10.1016/j.fct.2006.07.018

- Taylor, S.L., Baumert, J.L., Kruizinga, A.G., Remington, B.C., Crevel, R.W.R., Brooke-Taylor, S., Allen, K.J., Houben, G., 2014. Establishment of Reference Doses for residues of allergenic foods: Report of the VITAL Expert Panel. *Food Chem. Toxicol.* 63, 9–17. doi:10.1016/j.fct.2013.10.032
- van Rossum, C.T.M., Fransen, H., Verkaik-Kloosterman, J., Buurma-Rethans, E., Ocke, M., 2011. RIVM 2011, Dutch National Food Consumption Survey 2007-2010 : Diet of children and adults aged 7 to 69 years. RIVM-Rapport 350050006.
- Wainstein, B.K., Studdert, J., Ziegler, M., Ziegler, J.B., 2010. Prediction of anaphylaxis during peanut food challenge: usefulness of the peanut skin prick test (SPT) and specific IgE level. *Pediatr. Allergy Immunol.* 21, 603–611. doi:10.1111/j.1399-3038.2010.01063.x
- Zeiger, R.S., Sampson, H.A., Bock, S.A., Burks, A.W., Harden, K., Noone, S., Martin, D., Leung, S., Wilson, G., 1999. Soy allergy in infants and children with IgE-associated cow's milk allergy. *J. Pediatr.* 134, 614–622. doi:10.1016/S0022-3476(99)70249-0

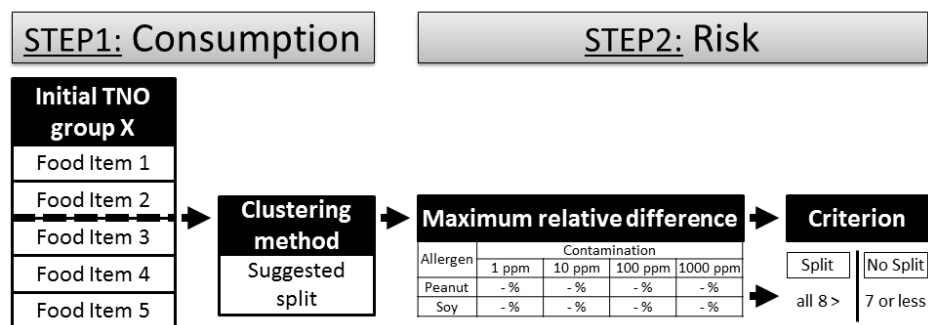


Figure 1: Procedure steps – summary

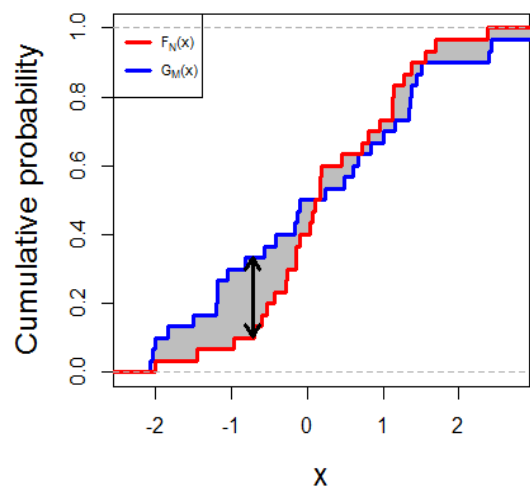


Figure 2: Illustrative example difference between KS and Cramér statistics (the arrow represents the KS statistic and the grey area the Cramér statistic)

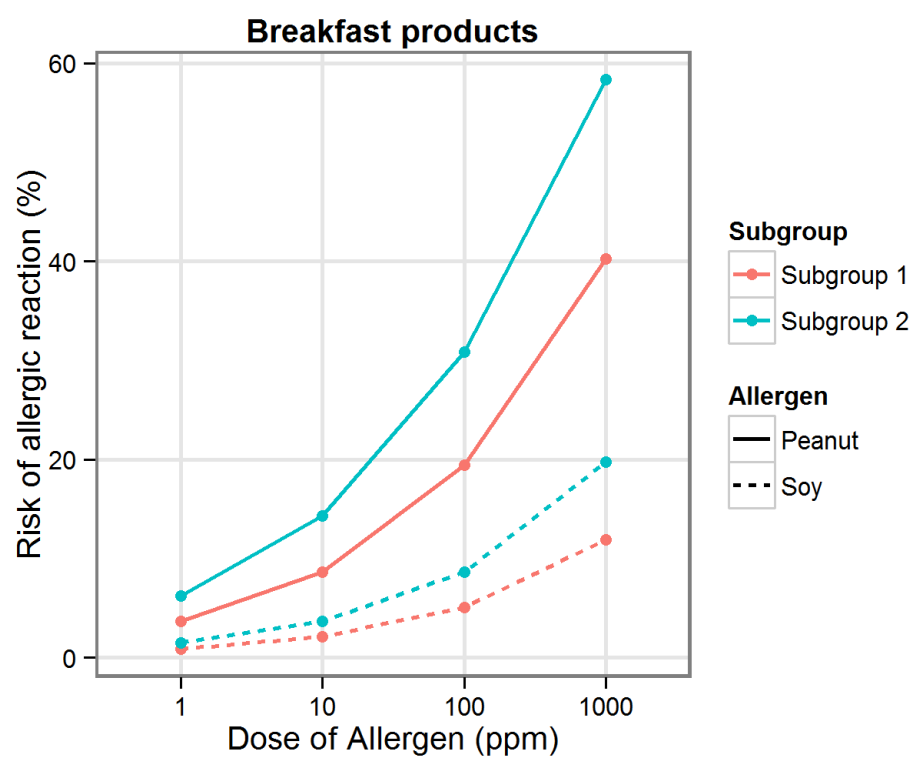


Figure 3: Risk of allergic reaction for two subgroups in the breakfast products group – breakfast product eaten unprocessed (1) and porridge (2)



Figure 4: Risk of allergic reaction for 2 subdivisions in the meat products group – mean consumption 30g (1) and 115g (2)

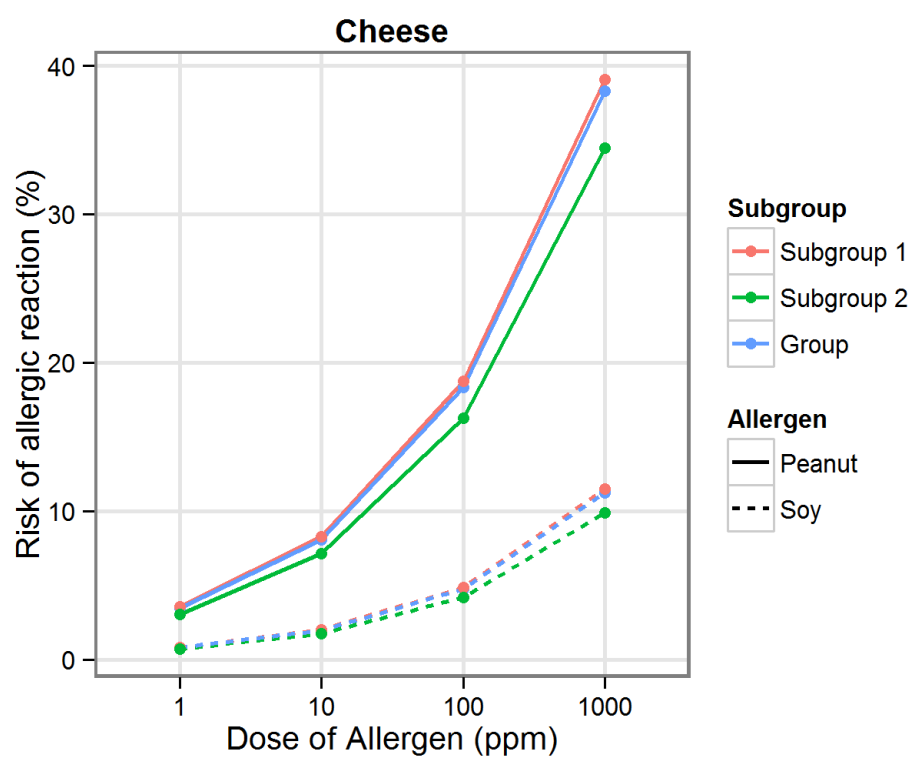


Figure 5: Risk of allergic reaction for the cheese group and for the 2 subdivisions

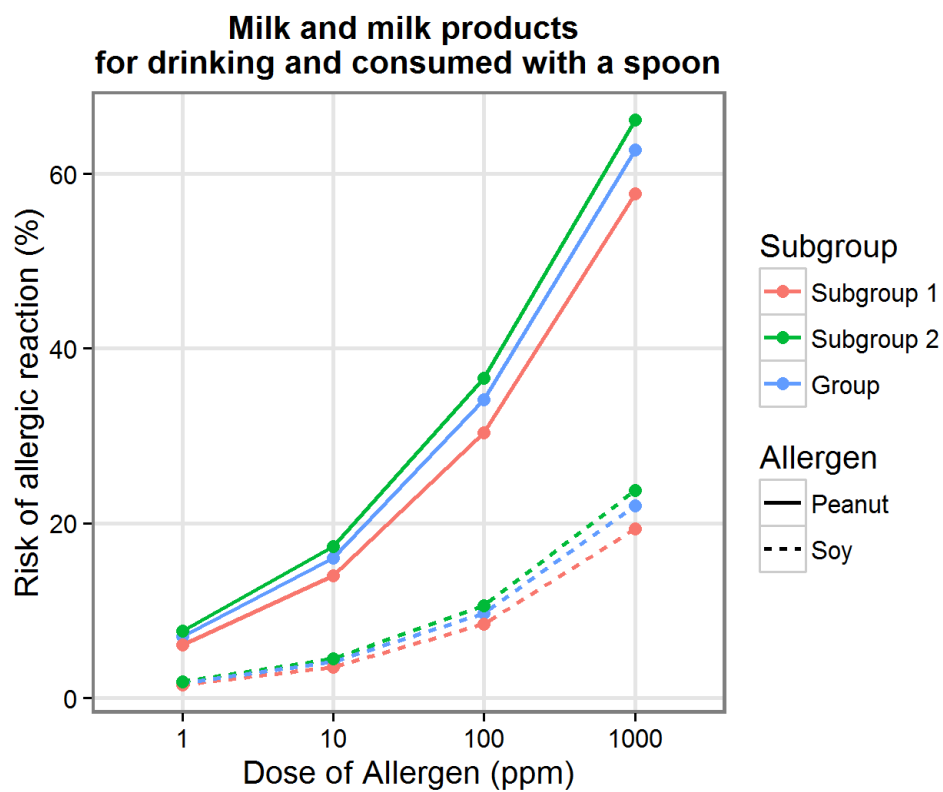


Figure 6: Risk of allergic reaction for the group milk and milk products for its 2 subdivisions

Table 1: Number of consumers and food items in each National Food Consumption Survey and number of groups based on the initial grouping method

Country	Number of consumers	Number of food items	Number of groups based on the initial grouping method
Netherlands	3819	496	42
France	2624	367	40
Denmark	2029	168	39

Table 2 Maximum relative difference in risk between breakfast product group and subgroups in Denmark

Allergen	Contamination			
	1 ppm	10 ppm	100 ppm	1000 ppm
Peanut	46.4%	44.6%	40.3%	31.6%
Soy	48.4%	48.8%	47.1%	44.2%

Table 3: number of subdivisions based on clustering and after the risk assessment analysis for Denmark, Netherlands and France

Number of subdivisions suggested by the clustering	No. of groups needing subdivision based on clustering			No. of groups needing subdivision based on risk assessment analysis		
	Denmark	Netherlands	France	Denmark	Netherlands	France
0	15	7	9	30	25	30
1	23	29	33	9	14	12
2	1	4	0	0	0	0

Table 4: Mean intake in the breakfast products group and subgroups

Group of interest	Mean intake	Standard deviation	Coefficient of variation
Overall group	75.0 g	100.0	133.4%
Subgroup 1	44.4 g	28.1	63.3%
Subgroup 2	168.0 g	163.2	97.2%

Table 5: Mean intake in the meat products group and subgroups

Group of interest	Mean intake	Standard deviation	Coefficient of variation
Overall group	57.9 g	59.1	102.1%
Subgroup 1	38.3 g	38.8	101.1%
Subgroup 2	111.9 g	70.7	63.2%

Table 6 Maximum relative difference in risk between cheese group and subgroups in Denmark

Allergen	Contamination			
	1 ppm	10 ppm	100 ppm	1000 ppm
Peanut	12.1%	11.9%	11.3%	10.1%
Soy	12.6%	12.2%	12.3%	12.0%

Table 7: Means intake in the cheese group and subgroups

Group of interest	Mean intake	Standard deviation	Coefficient of variation
Overall group	37.8 g	26.6	70.4%
Subgroup 1	40.6 g	25.0	61.6%
Subgroup 2	28.4 g	29.8	105.2%

Table 8 Maximum relative difference in risk between milk and milk products group and subgroups in Denmark

Allergen	Contamination			
	1 ppm	10 ppm	100 ppm	1000 ppm
Peanut	13.0%	12.6%	11.0%	8.1%
Soy	14.1%	13.7%	13.0%	12.1%

Table 9: Mean intake in the milk products group and subgroups

Group of interest	Mean intake	Standard deviation	Coefficient of variation
Overall group	260.3 g	180.2	69.2%
Subgroup 1	174.0 g	114.6	65.8%
Subgroup 2	324.6 g	192.9	59.4%

Appendix A: Food intake summary statistics per food group in Denmark (in g)

Final food group	Mean	SD	P50	P55	P60	P65	P70	P75	P80	P85	P90	P95	P975	P99	P100
Peanuts, nuts and dried fruits	34.1	29.4	30.0	30.0	30.0	30.0	30.0	40.0	40.0	60.0	60.0	90.0	120.0	150.0	330.0
Potato and other starch based chips (including salty sticks)	42.1	38.6	30.0	30.0	30.0	45.0	45.0	50.0	60.0	75.0	90.0	150.0	150.0	181.2	255.0
Fried/warm snacks	162.8	103.1	150.0	150.0	180.0	180.0	180.0	180.0	225.0	270.0	270.0	360.0	448.1	450.0	720.0
Meal replacements and meat imitates	40.1	18.7	32.0	32.0	32.0	38.4	51.2	64.0	64.0	64.0	64.0	64.0	64.0	64.0	64.0
Pancakes and waffles	151.5	104.6	100.0	150.0	150.0	150.0	150.0	200.0	200.0	250.0	300.0	335.0	400.0	500.0	800.0
Soups	395.9	169.8	375.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	750.0	750.0	750.0	1000.0
Small sweets	41.1	39.4	25.0	25.0	25.0	50.0	60.0	60.0	60.0	60.0	100.0	100.0	120.0	200.0	420.0
Sugar	21.4	12.9	16.0	16.0	22.0	23.0	23.0	24.0	32.0	32.0	33.0	46.0	56.8	68.0	120.0
Peanut butter	30.0	27.0	23.0	23.0	24.0	30.0	32.0	32.0	39.0	45.0	60.0	75.0	97.5	111.0	120.0
Chocolate and chocolate products	43.5	34.2	25.0	32.6	50.0	60.0	60.0	60.0	60.0	60.0	100.0	100.0	120.0	180.0	400.0
Sweet confectionary (jam, marmalade)	19.8	13.1	16.0	16.0	16.0	24.0	24.0	24.0	24.0	32.0	32.0	40.0	56.0	80.0	80.0
Cereal bars	27.2	24.2	25.0	25.0	25.0	25.0	25.0	25.0	25.0	32.4	45.0	50.0	50.4	92.6	250.0
Chewing gum	10.6	7.8	10.0	10.0	10.0	10.0	10.0	10.0	15.0	15.0	20.0	25.0	30.0	37.9	60.0
Mashed potato powder	151.3	87.9	139.0	139.0	139.0	139.0	139.0	212.0	222.0	222.0	222.0	340.0	436.2	444.0	680.0
Potato product (excl. powder)	221.3	120.9	225.0	225.0	225.0	250.0	300.0	300.0	300.0	375.0	375.0	434.0	464.0	530.0	1025.0
Vegetable oils and animal fat	21.0	16.7	16.0	16.0	20.0	24.0	24.0	30.0	32.0	32.0	45.0	49.0	64.0	80.0	150.0
Vegetable oils and animal fat	16.0	13.6	12.0	15.0	16.0	16.0	16.0	20.0	24.0	28.0	32.0	40.0	49.2	64.0	176.0
Sauces used as condiments and dessert sauces	26.1	22.4	15.0	21.0	30.0	30.0	30.0	30.0	30.0	45.0	45.0	60.0	75.0	116.1	450.0
Sauces , savory, chutneys and pickles	68.6	49.3	60.0	60.0	60.0	75.0	75.0	90.0	90.0	105.0	120.0	150.0	180.0	300.0	600.0
Cheese	37.8	26.6	32.0	32.0	32.0	40.0	46.0	46.0	48.0	64.0	69.0	90.6	96.0	134.4	300.0
Fish products – mean consumption 30 g ^a	30.8	25.4	23.0	23.0	23.0	30.0	32.0	40.0	40.0	46.0	60.0	80.0	102.5	127.5	160.0
Fish products - mean consumption 115 g ^b	122.9	110.6	100.0	101.0	101.0	110.0	126.0	150.0	162.0	200.1	202.0	303.0	324.0	458.5	1134.0
Meat products – mean consumption 65 g ^c	38.3	38.8	32.0	32.0	32.0	32.0	40.0	46.0	46.0	64.0	70.0	98.1	140.0	210.0	560.0
Meat products -mean consumption 105 g ^d	111.9	70.7	101.0	101.0	101.0	110.0	110.0	126.0	150.0	162.0	202.0	243.0	301.1	330.0	880.0
Crackers, crisp bread, rusk and toast	24.0	15.2	24.0	24.0	24.0	24.0	24.0	24.0	24.0	36.0	36.0	48.0	60.0	72.0	180.0
Bread, bread rolls and bread doughs	92.6	47.5	80.0	86.0	90.0	98.0	120.0	120.0	120.0	130.0	160.0	180.0	199.5	240.0	510.0
Herbs and spices mixes, bouillon cubes, yeast extract	18.1	32.4	10.0	10.0	10.0	15.0	20.0	20.0	20.0	20.0	20.0	35.5	65.5	200.0	400.0
Alcoholic drinks, alcohol ≤15%	272.8	149.9	280.0	280.0	280.0	280.0	280.0	280.0	420.0	420.0	420.0	560.0	633.5	700.0	1120.0

Alcoholic drinks, alcohol above 15%	61.8	50.6	50.0	60.0	60.0	60.0	60.0	60.0	90.0	90.0	100.0	150.0	180.0	240.0	699.0
Beer	672.8	624.4	330.0	495.0	660.0	660.0	660.0	660.0	990.0	990.0	1320.0	1980.0	2640.0	3300.0	5940.0
Drinks without alcohol (excl. syrup)	479.6	291.8	400.0	400.0	500.0	540.0	600.0	600.0	600.0	800.0	800.0	1000.0	1200.0	1600.0	5000.0
Cookies (biscuits)	37.6	25.5	30.0	36.0	36.0	36.0	47.5	48.0	48.0	60.0	72.0	96.0	120.0	120.0	192.0
Macaroons	14.9	13.1	7.0	14.0	14.0	14.0	14.0	14.0	21.0	22.4	28.0	44.8	56.0	57.1	70.0
Cakes (including pastry)	137.2	69.5	125.0	125.0	125.0	125.0	125.0	125.0	200.0	200.0	250.0	250.0	300.0	375.0	875.0
Breakfast products eaten unprocessed (e.g. müsli, oat and maize flakes)	44.4	28.1	46.0	46.0	46.0	46.0	47.0	64.0	64.0	64.0	92.0	92.0	117.0	117.0	234.0
Breakfast products, porridge	168.0	163.2	150.0	150.0	150.0	150.0	150.0	202.0	202.0	225.0	257.0	300.0	356.0	502.9	2850.0
Maize grain	56.2	45.7	38.0	38.0	38.0	60.0	64.0	64.0	64.0	114.1	115.0	119.6	215.0	215.0	345.0
Pasta, rice, couscous and other grains	159.7	102.0	125.0	139.0	139.0	148.0	208.0	208.0	222.0	222.0	278.0	363.0	416.0	444.0	726.0
Legumes ^e	109.5	70.4	82.0	106.0	106.0	106.0	106.0	130.8	139.0	139.0	215.0	230.4	282.7	363.9	444.0
Fruit and vegetables, processed	133.0	86.1	115.0	115.0	135.0	135.0	159.0	185.0	200.0	215.0	238.0	238.0	370.0	476.0	714.0
Eggs	42.1	28.7	40.0	40.0	46.0	55.0	55.0	55.0	55.0	55.0	80.0	110.0	110.0	110.0	165.0
Egg based dishes (Omelette plain, fried eggs)	68.2	38.0	50.0	50.0	50.0	79.1	100.0	100.0	100.0	100.0	100.0	100.0	150.0	200.0	300.0
Omellette with bacon	149.3	69.0	100.0	110.5	200.0	200.0	200.0	200.0	200.0	200.0	200.0	245.0	300.0	400.0	600.0
Sandwich and pizza	366.8	263.4	290.0	300.0	350.0	400.0	450.0	450.0	600.0	600.0	900.0	900.0	900.0	900.0	2000.0
Composite dishes ^f	204.9	122.1	165.0	165.0	208.0	237.0	237.0	237.0	248.0	330.0	358.0	450.0	474.0	716.0	1080.0
Cream and coffee milk	18.8	25.9	14.0	14.0	14.0	14.0	20.3	21.0	21.0	28.0	28.0	28.0	32.6	110.7	199.5
Ice cream	77.7	45.9	50.0	50.0	75.0	100.0	100.0	100.0	100.0	100.0	100.0	150.0	200.0	200.0	600.0
Milk and milk products for drinking and consumed with a spoon	260.3	180.2	200.0	200.0	200.0	200.0	250.0	400.0	400.0	400.0	400.0	600.0	600.0	800.0	3000.0

^a such as: prepared fish salad, fish fingers breaded, canned seafood

^b such as: fish balls

^c such as: bacon, cooked sausages, pat

^d such as: meat loaf, meat balls, meat burger

^e such as: dry seeds

^f such as: lasagne, goulash, fish gratin

Appendix B: Final Danish food groups and the overall FoodEx 2 groups included

Final food groups	FoodEx 2 Code	Foods included in final group (FoodEx 2 Name)
Peanut butter	A01BN	Peanut butter
Cheese	A02QE	Cheese
Cheese	A02QF	Fresh uncured cheese
Cheese	A02TQ	Cheese, danbo
Cheese	A031A	Processed cheese and spreads
Cream and coffee milk	A02ML	Cream, plain
Ice cream	A02QA	Ice cream, milk-based
Ice cream	A036J	Water-based ice creams
Milk and milk products for drinking and consumed with a spoon	A02LV	Cow milk
Milk and milk products for drinking and consumed with a spoon	A02LY	Cow milk, whole
Milk and milk products for drinking and consumed with a spoon	A02LZ	Cow milk, semi skimmed (half fat)
Milk and milk products for drinking and consumed with a spoon	A02MA	Cow milk, skimmed (low fat)
Milk and milk products for drinking and consumed with a spoon	A02MP	Flavoured milks
Milk and milk products for drinking and consumed with a spoon	A02NG	Yoghurt, cow milk, plain
Milk and milk products for drinking and consumed with a spoon	A02NH	Yoghurt, cow milk, flavoured
Milk and milk products for drinking and consumed with a spoon	A02NQ	Yoghurt drinks, sweetened and/or flavoured
Milk and milk products for drinking and consumed with a spoon	A02QD	Milkshakes
Milk and milk products for drinking and consumed with a spoon	A03TJ	Soya drink
Milk and milk products for drinking and consumed with a spoon	A03TM	Rice drink
Milk and milk products for drinking and consumed with a spoon	A041E	Rice pudding
Milk and milk products for drinking and consumed with a spoon	A04NS	Other desserts spoonable
Peanuts, nuts and dried fruits	A014C	Tree nuts
Peanuts, nuts and dried fruits	A014K	Coconut
Peanuts, nuts and dried fruits	A015F	Oilseeds
Peanuts, nuts and dried fruits	A015H	Peanut
Peanuts, nuts and dried fruits	A01MA	Dried fruit
Peanuts, nuts and dried fruits	A01MB	Dried prunes
Peanuts, nuts and dried fruits	A01MD	Dried apricots
Peanuts, nuts and dried fruits	A01MF	Dried dates
Peanuts, nuts and dried fruits	A01MG	Dried figs
Peanuts, nuts and dried fruits	A01QF	Mixed dried fruits
Potato and other starch based chips (including salty sticks)	A00DC	Popcorn (maize, popped)
Potato and other starch based chips (including salty sticks)	A011L	Potato crisps (chips)
Fried/warm snacks	A040D	Sausage roll
Fried/warm snacks	A040F	Spring rolls
Meal replacements and meat imitates	A03RS	Food for weight reduction
Pancakes and waffles	A00CL	Pancakes
Soups	A041S	Other vegetable soup
Soups	A041V	Meat and vegetable soup
Soups	A042A	Fruit soup
Small sweets	A034V	Sweet confectionery
Small sweets	A034X	Candies
Small sweets	A034Y	Marzipan
Small sweets	A035H	Foamed sugar products (marshmallows)
Small sweets	A035J	Liquorice candies
Small sweets	A035K	Gum drops
Sugar	A032H	Sucrose (common sugar)
Sugar	A04PA	Sugar and other sweetening ingredients

		(excluding intensive sweeteners)
Chocolate and chocolate products	A034F	Chocolate
Sweet confectionary (jam, marmalade)	A01MM	Jam
Cereal bars	A00EY	Cereal bars
Cereal bars	A00FH	Mixed cereal-based snacks
Chewing gum	A035M	Chewing gum
Mashed potato powder	A011E	Mashed potato powder
Potato product (excl. powder)	A011P	Potato boiled
Potato product (excl. powder)	A011R	Potato baked
Potato product (excl. powder)	A03VD	Potato based dishes
Potato product (excl. powder)	A03VG	Prepared potato salad
Potato product (excl. powder)	A0BYV	French fries from cut potato
Vegetable oils and animal fat	A036P	Olive oils
Vegetable oils and animal fat	A036V	Rape seed oil, edible
Vegetable oils and animal fat	A037V	Pork lard
Vegetable oils and animal fat	A039E	Blended margarine
Vegetable oils and animal fat	A04SD	Blended fat and oils
Vegetable oils and animal fat	A037V	Pork lard
Butter/halvarine/margarine	A039C	Butter
Butter/halvarine/margarine	A039D	Traditional margarine
Sauces , savory, chutneys and pickles	A043V	Savoury sauces
Sauces , savory, chutneys and pickles	A044E	Vegetables-based cooked sauce
Sauces used as condiments and dessert sauces	A044F	Table-top condiments
Sauces used as condiments and dessert sauces	A044R	Soy sauce
Sauces used as condiments and dessert sauces	A044V	Pesto
Sauces used as condiments and dessert sauces	A044X	Mayonnaise
Sauces used as condiments and dessert sauces	A045J	Mixed and other not listed condiments
Sauces used as condiments and dessert sauces	A045K	Salad dressing
Sauces used as condiments and dessert sauces	A045L	Salad dressing, low fat
Sauces used as condiments and dessert sauces	A045N	Tartar sauce
Fish products - mean 30 g	A02KC	Fish fingers, breaded
Fish products - mean 30 g	A03XP	Prepared fish salad
Fish products - mean 30 g	A0BZ5	Canned seafood
Fish products - mean 110 g	A02KD	Fish balls
Meat products - mean 65 g	A022X	Bacon
Meat products - mean 65 g	A023G	Cooked cured meat
Meat products - mean 65 g	A025J	Cooked sausages
Meat products - mean 65 g	A025Z	Head cheese
Meat products - mean 65 g	A026R	Pate, pork liver
Meat products - mean 65 g	A04NG	Sausages and other comminuted meat
Meat products -mean 105 g	A026D	Wiener sausage
Meat products -mean 105 g	A03VV	Meat based dishes
Meat products -mean 105 g	A03XA	Meat loaf
Meat products -mean 105 g	A03XD	Meat loaf with cheese, vegetables or other
Meat products -mean 105 g	A03XF	Meat burger (no sandwich)
Meat products -mean 105 g	A03XG	Meat balls
Crackers, crisp bread, rusk and toast	A006D	Wheat crisp bread
Bread, bread rolls and bread doughs	A004V	Bread and similar products
Bread, bread rolls and bread doughs	A004Y	Wheat bread and rolls, white (refined flour)
Bread, bread rolls and bread doughs	A005E	Wheat bread and rolls, brown or wholemeal
Bread, bread rolls and bread doughs	A005H	Rye bread, wholemeal
Bread, bread rolls and bread doughs	A006S	Pita bread
Bread, bread rolls and bread doughs	A006V	Tortilla
Herbs and spices mixes, bouillon cubes, yeast extract	A01AM	Caper buds

Herbs and spices mixes, bouillon cubes, yeast extract	A042Y	Seasoning mixes
Herbs and spices mixes, bouillon cubes, yeast extract	A043H	Stock cube, beef flavour
Alcoholic drinks, alcohol ≤15%	A03MV	Wine, white
Alcoholic drinks, alcohol ≤15%	A03MX	Wine, red
Alcoholic drinks, alcohol ≤15%	A03ND	Cider
Alcoholic drinks, alcohol above 15%	A03NG	Fortified and liqueur wines
Alcoholic drinks, alcohol above 15%	A03PD	Unsweetened spirits
Alcoholic drinks, alcohol above 15%	A04QG	Other spirits
Beer	A03MD	Beer, strong
Beer	A03ME	Beer, regular
Beer	A03MF	Beer, light
Drinks without alcohol (excl. syrup)	A03KC	Coffee (average strenght) beverage
Drinks without alcohol (excl. syrup)	A03KG	Coffee with milk or cream
Drinks without alcohol (excl. syrup)	A03LB	Tea beverages
Drinks without alcohol (excl. syrup)	A03KH	Coffee drink, cappuccino
Drinks without alcohol (excl. syrup)	A03LG	Herbal and other non-tea infusions
Drinks without alcohol (excl. syrup)	A039L	Fruit juices
Drinks without alcohol (excl. syrup)	A03CH	Vegetable juices, ready to drink
Drinks without alcohol (excl. syrup)	A04PX	Bottled water
Drinks without alcohol (excl. syrup)	A03DZ	Soft drinks
Drinks without alcohol (excl. syrup)	A03FT	Diet soft drinks
Cookies (biscuits)	A009V	Biscuits (sweet and semi-sweet)
Cookies (biscuits)	A009X	Biscuits, sweet, plain
Cookies (biscuits)	A00AE	Biscuit, filled (with inclusions, filling or coating)
Macaroons	A00CN	Macaroons
Cakes (including pastry)	A00AV	Cream cake
Cakes (including pastry)	A00BZ	Fruit pie-tarts
Cakes (including pastry)	A00CC	Puff pastry
Breakfast products eaten unprocessed (e.g. müsli, oat and maize flakes)	A00DH	Oat rolled grains
Breakfast products eaten unprocessed (e.g. müsli, oat and maize flakes)	A00EJ	Muesli and similar
Breakfast products eaten unprocessed (e.g. müsli, oat and maize flakes)	A04QY	Other processed and mixed breakfast cereals
Breakfast products, porridge	A00EN	Porridge
Breakfast products, porridge	A00ES	Rye porridge
Maize grain	A000T	Maize grain
Pasta, rice, couscous and other grains	A003E	Rice grain, parboiled
Pasta, rice, couscous and other grains	A004G	Bulgur
Pasta, rice, couscous and other grains	A007D	Pasta and similar products
Pasta, rice, couscous and other grains	A040Z	Rice based dishes
Legumes	A012R	Pulses (dry)
Fruit and vegetables, processed	A01QG	Fruit salad
Fruit and vegetables, processed	A03YE	Mixed vegetables, boiled
Fruit and vegetables, processed	A03YF	Vegetables, gratinated
Fruit and vegetables, processed	A042B	Salads
Fruit and vegetables, processed	A042F	Greek salad
Eggs	A031S	Hen egg yolk
Eggs	A032B	Boiled eggs
Egg based dishes (Omelette plain, fried eggs)	A032C	Fried eggs
Egg based dishes (Omelette plain, fried eggs)	A03YN	Omelette, plain
Omelette with bacon	A03YP	Omelette with bacon
Sandwich and pizza	A03ZA	Sandwich with cheese topping/filling
Sandwich and pizza	A03ZG	Sandwich with meat and vegetable topping/filling
Sandwich and pizza	A03ZK	Hot dog with bread
Sandwich and pizza	A03ZL	Hamburger with bread
Sandwich and pizza	A040B	Pizza with cheese, meat, mushrooms, and vegetables

Composite dishes	A03VK	Potatoes, meat, and vegetables meal
Composite dishes	A03VS	Beans and vegetables meal
Composite dishes	A03VX	Goulash
Composite dishes	A03VY	Meat stew
Composite dishes	A03XJ	Fish and seafood based dishes
Composite dishes	A03XL	Fish gratin
Composite dishes	A03XM	Seafood-based meals
Composite dishes	A03XS	Fish and vegetables meal
Composite dishes	A03XZ	Vegetable casserole
Composite dishes	A03YA	Veggie pot pie
Composite dishes	A03YM	Quiche
Composite dishes	A040P	Lasagna
Composite dishes	A041J	Rice, meat, and vegetables meal